Knowledge and Productivity in the Pharmaceutical Industry

Alexandra Stone and Kenneth Flamm

Lyndon B. Johnson School of Public Affairs, University of Texas at Austin

P.O. Box Y

Austin, TX USA 78713-8925

***Abstra ct*—Pharmaceutical firms perform R&D to develop and maintain core knowledge and capabilities. They use foundational knowledge to identify and evaluate relevant scientific advances and integrate those advances with existing knowledge and capabilities throughout the innovation pipeline. If core scientific knowledge and technical capabilities are used in this manner to increase innovative productivity by decreasing the use of other inputs, conventional measures of productivity based solely on outputs, per unit time, are likely to be inaccurate. We analyze productivity by examining development time and controlling for knowledge inputs. This approach measures the effect of core knowledge generated through pre-clinical laboratory R&D on development time in later stages of clinical development. We use patents and citations to model the use of technical knowledge generated through pre-clinical laboratory research in subsequent clinical development activities. Our results indicate that internally generated knowledge and patents, which builds on the existing scientific base are associated with shorter development times.**

***Keywords*— pharmaceutical industry; innovation; R&D; patents; absorptive capacity**

I. Introduction

Pharmaceutical innovation is driven by scientific knowledge [1]-[4]. Research and Development (R&D) yields scientific advances that facilitate the discovery of new drug candidates and generates technical knowledge and capabilities that enable drug development. Accordingly, pharmaceutical firms perform continuous R&D to maintain an innovation capacity. To spur drug development, U.S. Federal government and pharmaceutical companies doubled R&D spending in the late 1990s [1], [3]. Stakeholders from the public and private sectors expected to observe more productive development pipelines filled with greater numbers of drug candidates leading to the successful development of greater numbers of novel drugs. However, the anticipated results of increased R&D spending have not been realized. U.S. pharmaceutical companies received approval from the US Food and Drug Administration (FDA) for fewer drug candidates in the early 2000s, despite increases in R&D spending, than in the early 1990s [5]. Measures of returns to R&D spending in terms of the number of FDA approved drugs indicate stagnating productivity [1], [2], [5].

Productivity is a measure of outputs as a function of total inputs to development and production activities. In the pharmaceutical industry, R&D productivity is often measured using the number of drugs approved or the number of drug candidates successfully completing a phase of clinical development. Key inputs into clinical drug development include R&D expenditures, number of employees, and knowledge inputs measured as number of patents. Such productivity measures are based on the premise that knowledge is a critical input to innovation and productivity in the pharmaceutical industry. A firm's ability to develop a new drug depends on its absorptive capacity, which is defined by foundational knowledge and technical capabilities developed through continuous R&D [6]. Core knowledge and capabilities that complement new knowledge enables the effective use of scientific advances [6], [7]. Measures of productivity in the pharmaceutical industry should capture the effective use of knowledge to enable innovation.

The knowledge and capabilities generated through core R&D are often used across multiple development activities, often involving multiple drug candidates [1]. Firms use core knowledge to identify and evaluate relevant scientific advances and integrate those advances with existing knowledge and capabilities to increase productivity [6]. If core scientific knowledge and technical capabilities are used in this manner to increase innovative productivity by decreasing the use of other inputs, conventional measures of productivity based solely on outputs, per unit time, are likely to be inaccurate. Public policies affecting pharmaceutical innovation based on incorrect premises are likely to be ineffective.

In this paper, we analyze productivity by examining development time and controlling for knowledge inputs. This approach measures the effect of core knowledge generated through pre-clinical laboratory R&D on development time in later stages of clinical development. We use patents and citations to model the use of technical knowledge generated through pre-clinical laboratory research in subsequent clinical development activities. We use the assignee of the cited patents to differentiate between knowledge generated through internal R&D and knowledge incorporated from external sources. The novel approach presented here focuses on the improvements in development time, controlling for other inputs, derived

This research was supported by NSF award number 0830389: Modeling Innovation Chains Using Case-Based Econometrics: Nano-electronics and Biotechnology Applications.

from the technical capabilities embodied in firm-specific knowledge.

II. Hypotheses

We model the effect of the firm's use of internally and externally generated knowledge on drug development time, from date of oldest patent grant to date of FDA approval. Patents are one measure of novel knowledge and technological capabilities, which have the potential to be developed into a marketable product [2], [4], [8], [9]. As in previous research, we use patents as proxies for technical knowledge incorporated into development activities. Previous studies provide empirical evidence that the knowledge stocks described in patents are associated with improved performance.[1] We build on previous research by examining the use of scientific and technical knowledge to increase productivity by decreasing development time.

We hypothesize that development time is negatively associated with the number of patents cited in the drug approval application. Development time is likely to be shorter when firms integrate more patented knowledge and capabilities into a development pipeline than when firms use less patented knowledge. We also hypothesize that the repeated incorporation of scientific and technological advances throughout the development process is negatively associated with development time. Development time is likely to decrease as the time between the issue of the oldest referenced patent and issue of the newest referenced patent increases.

Firms combine internally and externally generated knowledge in performing innovation activities. Scientific and technical advances that are related to the firm's core knowledge are more easily integrated into innovation activities than advances that are not related [6], [10]. Since knowledge and capabilities are generated in an evolutionary manner, the firm's core knowledge is more likely to complement internally generated advances than externally generated advances [4], [6]. Firms are more likely to have to conduct additional foundational research to enable the use of externally generated knowledge.

We hypothesize that development time is negatively associated with the use of internally generated knowledge. Development time is likely to be shorter when firms integrate more internally generated knowledge and capabilities into a development pipeline than when more externally generated knowledge is used.

A patent's citations to prior art identify the knowledge that was utilized in the conception and development of the patented invention [8]. A greater number of citations to prior patents indicates a stronger connection to existing scientific knowledge and technical capabilities [9], [12], [13]. In using such patents, firms build on a more robust scientific and technological base.

We hypothesize that development time is likely to be negatively associated with the strength of a patent's connection to the scientific and technological base. Development time is likely to be shorter when firms utilize patents that are strongly connected to the scientific base than when patents that are less strongly connected to the scientific base are used.

More recent scientific advances are associated with technical uncertainty. Firms are likely to have to do additional research in order to apply patents that embody newer technical advances effectively [6], [10]. Such additional research is likely to result in longer development times.

We hypothesize that development time is likely to be negatively associated with the age of the knowledge cited by the patent as prior art. Development time is likely to be shorter when firms integrate patented technical advances that build on older knowledge and technology than when they integrate advances based on more recent advances.

Patents are novel inventions with the potential to be developed into a marketable product. In more novel patents, knowledge and capabilities are used in a broader range of applications that may be developed commercially [9]. In previous research, the number of novel components, delineated in the patent's claims, was positively correlated with the likelihood that a patent would be commercialized [12]. Broader patents, which make more claims, are associated with greater commercial opportunity. We build on this research by examining the association between the breadth of novel claims made by a patent and the time to develop the patented technology into an FDA approved drug. Broad patents with many claims involve more novel knowledge which may be more technically difficult to utilize. Firms may have to perform additional research to utilize the novel components of patents with more claims [6], [10].

We hypothesize that development time is positively associated with the number of claims. Since the patent's claims delineate the novel components, development time is likely to increase when firms utilize patents with more novel components.

Pharmaceutical firms are likely to utilize knowledge generated by different types of entities, including government laboratories, universities, and other firms, from around the world [1]. As set forth in the organizational learning theories that describe the firm's absorptive capacity, a firm's ability to utilize knowledge depends on the relatedness of the new knowledge to existing knowledge stocks [6], [10]. A firm's ability to utilize new externally generated knowledge is not likely to be determined by the national origin of the entity that generated the knowledge.

We hypothesize that development time is not affected by greater reliance on knowledge generated by non-U.S. firms and universities. Among firms using externally generated knowledge, development times for those using knowledge generated by U.S. entities will be similar to those using knowledge generated by non-U.S. entities.

III. Data

[1] Firm performance has been measured as sales, market value, profitability, and number of new products [2, 4, 11, 12].

We have created a new data set that links FDA approved drugs to drug-related patents.2 The analysis is limited to the final stages of the innovation pipeline, between initial patent grant and FDA approval, which include clinical trials and FDA review. The patents associated with each new drug reflect the knowledge inputs used in these innovation activities. These innovation activities are performed by the firm that is granted FDA approval (the focal firm). The sample consists of 343 new drugs developed by 194 firms and approved between 1985 and 2008. The drug approval applications included in the sample cite 252 patents. We obtained data on the referenced patents from the National Bureau of Economic Analysis Patent Database [8]. The date of issue for the patent used to develop a drug is included in the sample, and demarcates the beginning of this phase in the innovation pipeline.3 More than one firm may use the same patent to produce a new drug requiring FDA approval. Each firm's new drug application and reference to a patent reflect its first use of patented knowledge on a given development project, but not necessarily the very first use of that knowledge. We construct development time as the time elapsed from the issue date for the oldest patent referenced in the drug application to the FDA approval date.

IV. Analytic Approach

A. Analytic Method

We use Poisson regression models to predict development time, measured in months, from patent issue to FDA approval.[4] Poisson regression models assume that the dependent variable takes on non-negative integer values, with a mean equal to the exponential of a linear function of explanatory variables [14]. With a Poisson distribution, variance is equal to a random variable's mean (equidispersion). The Poisson model will produce consistent estimates of the effects of the explanatory variables, even if the distribution of the explanatory variable is not equidispersed, as long as the expectation of the dependent variable conditional on the explanatory variables is correctly specified [15].[5] Robust standard errors can be calculated to provide correct parameter variance estimates if the actual data generating process deviates from the equidispersion process. We therefore use Poisson quasi-maximum likelihood methods, couples with robust standard errors, to estimate the effects of explanatory variables on development time in a fashion that minimizes distributional assumptions.

B. Explanatory Variables

All of the drug applications used in this sample cite at least one patent that was used in the development of that drug. We use the number of patents cited in an application as a general measure of the amount of knowledge used in innovation activities. We use the time lag from the date on which the oldest patent was issued to the date on which the most recent patent was issued to measure the repeated incorporation of new scientific knowledge into development processes. Larger lags are indicative of the repeated use of new advances embodied in patents. Some of the patents in the drug approval application are assigned to the focal firm, while others are assigned to other entities. We use the proportion of patents cited in the approval application that were assigned to the focal firm as a measure of the internal knowledge used to develop the drug. These patent variables reflect a firm's general use of scientific knowledge and its reliance on internally developed knowledge.

We use several patent characteristics to characterize the knowledge inputs used in innovation activities. We use the number of claims made by a patent to measure the number of novel components in a patent and the breadth of knowledge and capabilities embodied in the patent. We use the number of citations made to patents in the prior art as a measure of the connection between the knowledge embodied in the patent and existing scientific knowledge. We use a patent's mean backward citation lag to measure the age of the knowledge used in early stage R&D that led to the patented invention. The citation lag to the prior art reflects the maturity of the knowledge and technologies used in pre-clinical laboratory development activities. These citation-based measures delineate the relationship between the knowledge embodied in the patent and the existing scientific and technological base.

We use a series of dummy variables to distinguish between focal patents assigned to U.S. and foreign entities.[6] Assignee type one refers to “unassigned” patents that were issued to the inventor. Information on the nationality of the assignee is not available, since the patent is not formally assigned. Assignee type two is comprised of U.S. non-government entities, including firms and universities. Assignee type three includes foreign non-government entities.[7] The coefficients for each assignee type measure the effect on development time relative to the use of patents assigned to U.S. non-government entities (type two).

We control for some firm-specific characteristics that are likely to affect productivity. To control for firm size, we use the number of employees in the year before the firm receives FDA approval for the drug. We use the number of new drugs for which the firm has received FDA approval to control for the firm's experience in clinical development and managing the FDA approval process.

---

2 We obtained data on FDA approved drugs from the FDA Orange Book, which can be accessed at www.accessdata.fda.gov/scripts/cder/ob/default.cfm

3 We identified a firm's first use of a patent through reference to the patent in the FDA drug approval application.

4 Although the patent application date may be a more relevant indicator of commercial and technical potential, the exact application date is not included in the NBER patent database.

5 This robustness property requires that the true distribution be in the linear exponential family, which includes among its members the Poisson, negative binomial, Bernoulli, exponential, gamma, normal, and inverse Gaussian distributions.

6 The assignee variable refers to the assignee of the focal patent. It does not indicate whether the patent was assigned to the company who developed the drug and received FDA approval.

7 This assignee classification system is based on the USPTO assignee classification system [8]. It includes government agencies as assignees. However, none of the patents included in the sample were assigned to government agencies.

We control for the implementation of new regulatory policies implemented under the Prescription Drug User Fee Act (PDUFA). The Act was intended to improve the approval process by authorizing the FDA to collect user fees that are used to augment resources used in reviewing drug approval applications. The PDUFA is a dummy variable equal to one for drugs approved after the implementation act and zero for those approved before the act was implemented. Although regulatory policies are not likely to affect R&D strategies to develop knowledge stocks, previous research indicates that PDUFA has affected drug approval time [5].

V. Results

We obtained data on a sample of 343 FDA approved drugs that cited patents in their drug approval applications.[8] Development time ranged from one to 211 months, with an average of 77.6337 months. We present descriptive statistics for all explanatory variables in Table I and a histogram illustrating the distribution of development times in Fig. 1.

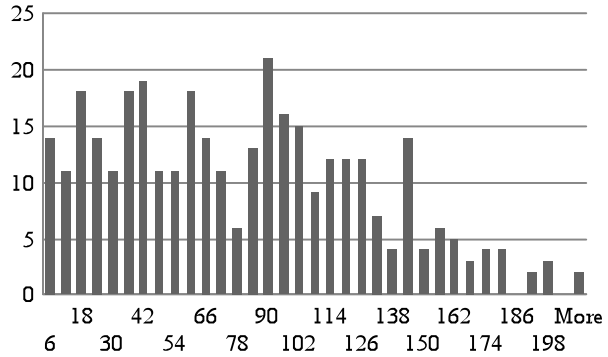

Fig. 1. Histogram depicting the distribution of drug development times

TABLE I
Descriptive Statistics

	Mean	Standard Deviation	Minimum	Maximum
Development time	77.6337	48.8308	1	211.0000
Patents per FDA application	3.0029	2.4020	1	18.0000
Time from oldest cited patent to most recent cited patent	25.4800	36.0600	0	176
Proportion of patents assigned to the firm	0.5369	0.4625	0	1
Total number or approved drugs	4.1000	5.7600	0	54
Number of claims	17.0000	13.5200	1	113
Number of citations made	8.3300	9.3600	1	63
Citation lag from focal patent to cited prior art	10.2400	6.2700	1	48.25
Employees (thousands)	32.1300	36.6195	0.0080	151.9
Prescription Drug User Fee Act	0.9700	0.1682	0	1
Assignee type 1	0.0436	0.2045	0	1
Assignee type 2	0.5639	0.4966	0	1
Assignee type 3	0.3866	0.4877	0	1
Assignee type 4	0.0058	0.0761	0	1

In Table II, we present the results of the Poisson regression model and robust standard errors used to estimate the effects of the use of knowledge inputs, embodied in patents, on development time. The Wald test statistic for the overall model is significant. The results indicate that the variables included in the model have a joint effect on development time, and thus we reject the null hypothesis of no joint effect. One measure of goodness of fit for this model is the correlation of predicted development times with actual development times [15]. The correlation is 0.1402 and statistically significant. Although the correlation is low, it is within the typical ranges for R2-like measures for models using micro data with substantial idiosyncratic variation across subjects. Thus, the knowledge inputs embodied in patents affect productivity, measured as development time.

A number of patent characteristics have a statistically significant effect on development time (Table II). The regression coefficients are semi-elasticities that measure the proportional change in development time from a unit change in the covariate. We present the elasticities of development time associated with each explanatory variable in Table 3. Although the number of patents referenced in the FDA approval application was not statistically significant, the proportion of patents assigned to the focal firm and the lag from oldest to newest patent have statistically significant effects. In accordance with our hypothesis, a 1% increase in the proportion of referenced patents assigned to the focal firm is a associated with a 0.09% decrease in development time. The use of internally generated knowledge is associated with shorter development times. Contrary to our hypothesis, a 1% increase in lag from the issue date of the oldest patent to the issue date of the newest patent is associated with a 0.06% increase in development time. The number of citations made by the focal patent has a statistically significant negative effect, while the time from the application date of the focal patent to the issue date of the cited patent is not statistically significant. As we hypothesized, a 1% increase in the number of citations made is associated with a 0.07% decrease in development time. These results suggest that a stronger connection to existing knowledge decreases development time regardless of the age of the cited knowledge. The number of claims included in the patent had a statistically significant positive effect. In accordance with our hypothesis, a 1% increase in the number of claims made by a patent is associated with a 0.06% increase in development time. The observed effect suggests that the use of patents with greater numbers of novel elements increases development time. Thus, the observed results provide preliminary evidence that knowledge inputs can

[8] The sample only includes FDA drug approval applications that cite patents that were issued prior to drug approval. Firms may receive patents for drug-related inventions issued after approval.

be used to increase productivity by decreasing development time.

Firms utilize knowledge from different sources. Of the 343 FDA approved drugs included in this sample, 100 cited patents that are assigned to non-U.S. government entities. The dummy variable for patents assigned to foreign entities is not statistically significant. This lack of significance indicates that the national origin of the patented knowledge does not affect development time. These results provide evidence of the increasing international dynamics of pharmaceutical innovation and the diminishing importance of a national innovation ecosystem.

The Prescription Drug User Fee Act had a statistically significant effect on drug development time. However, contrary to our hypothesis and previous research, development time increased after implementation of the Act. The observed increase indicates that changes in the bureaucratic process increase development and approval time. It is important to note that the mean approval year was 1998 and that most drugs were approved after the act was implemented. In future research, we will explore the effects of policies to promote pharmaceutical innovation and the impacts of regulatory policies.

TABLE II
REGRESSION RESULTS

	Coefficients (standard error) * statistically significant at 0.1 level
Wald Chi-square statistic	120.8
Constant	2.6772 * (0.2382)
Patents per FDA application	- 0.0093 (0.0146)
Time from oldest cited patent to most recent cited patent	0.0025 * (0.0008)
Proportion of patents assigned to the firm	- 0.1736 * (0.0703)
Total number of FDA approved drugs	0.0018 (0.0049)
Number of claims	0.0035 * (0.0020)
Number of citations made to prior art	- 0.0086 * (0.0041)
Citation lag from focal patent to cited prior art	- 0.0077 (0.0049)
Number of employees	- 0.0012 (0.0009)
Prescription Drug User Fee Act	1.8174 * (0.2197)
Assignee type 1	0.1868 (0.1199)
Assignee type 3	0.0914 (0.0676)
Assignee type 4	-0.05783 (0.3866)
Correlation between predicted development time and observed development time	0.14

VI. CONCLUSION

We have presented empirical evidence that scientific knowledge and technical capabilities may be used to enhance R&D productivity in the pharmaceutical industry by decreasing development time. According to the organizational learning theory of absorptive capacity, firms use existing knowledge and technical capabilities to integrate new knowledge into development activities.

Our results indicate that the utilization of patents that build on existing knowledge, as indicated by the number of citations made to the prior art, leads to shorter development times. Moreover, knowledge generated through internal R&D activities, which we identified as patents assigned to the firm performing clinical development activities, is associated with shorter development time. Internally generated knowledge constitutes a knowledge base that firms use to integrate new knowledge that enhances R&D productivity. Thus, R&D that builds on existing knowledge and capabilities increases productivity by reducing development time.

TABLE III
ELASTICITY OF DEVELOPMENT TIME

Variables	Elasticity
Patents per FDA application	-0.0279
Time from oldest cited patent to most recent cited patent	0.0649
Proportion of patents assigned to the firm	-0.0932
Total number of FDA approved drugs	0.0072
Number of claims	0.0602
Number of citations made to prior art	-0.0719
Citation lag from focal patent to cited prior art	-0.0785
Employees	-0.0380
Prescription Drug User Fee Act	1.76
Assignee type 1	0.0082
Assignee type 3	0.0352
Assignee type 4	-0.0003

Scientific advances lead to the development of novel drugs [1]. However, new knowledge does not necessarily enhance productivity by decreasing development time. Prior research has used the number of claims made by a patent as a quantitative indicator of the pioneering nature of a patented invention that was associated with an increased likelihood that the invention would be commercialized [12]. Our results show that development time is negatively associated with the “pioneering” nature of an invention. The utilization of patents with more novel components, identified as patent claims, increases development time. The positive effect of the time lag from the date on which the oldest patent was issued to the

date on which the most recent patent was issued indicates that development time increases with the ongoing incorporation of new scientific knowledge into development activities. Our results indicate that the use of more "novel" scientific knowledge increases development time. According to organizational learning theory, firms conduct additional research to enable the effective use of "novel" knowledge that is not related to their existing knowledge [6], [10]. Thus, the returns to basic scientific advances are likely to be realized further in the future because of the complexity of utilizing new knowledge in ongoing development activities.

We examined the effect of the Prescription Drug User Fee Act, which was intended to accelerate innovation by providing the FDA additional resources to review approval applications in a timely manner. Although some have found that PDUFA has reduced mean development time, our results indicate that development time increased after implementation of PDUFA [5]. It is possible that PDUFA has decreased development time, but that the decrease has been outweighed by other time-variant factors that have increased development time. Nonetheless, the positive effect of PDUFA indicates that development time has increased.

Our results indicate that firms use core knowledge to enhance R&D productivity. Accordingly, measures of productivity that do not account for the role of R&D in sustaining core competencies will be imprecise. Policies intended to promote innovation in the pharmaceutical industry should reflect the enabling role of knowledge in development activities. In future research, we will examine the effects of innovation policies on the use of knowledge to improve productivity.

ACKNOWLEDGMENT

The authors thank Karthick Santhanam for his research support.

REFERENCES

[1] Cameron, A.C. and P.K. Trivedi, *Microeconometrics Using Stata*. 2009, College Station Texas: Stata Press. 692.

[2] Cockburn, I.M., *The Changing Structure of the Pharmaceutical Industry.* Health Affairs, 2004. **23**(1): p. 10 - 22.

[3] Cohen, W.M. and D.A. Levinthal, *Absorptive Capacity: a New Perspective on Learning and Innovaiton* Administrative Science Quarterly, 1990. **35**.

[4] DeCarolis, D.M. and D.L. Deeds, *The Impact of Stocks and Flows of Organizational Knowledge on Firm Performance: an Empirical Investigation of the Biotechnology Industry.* Strategic Management Journal, 1999. **20**: p. 953 - 968.

[5] DiMasi, J. and H.G. Grabowski, *The Cost of Biopharmaceutical R&D: Is Biotech Different?* Management and Decision Economics, 2007. **28**: p. 469 - 479.

[6] Hall, B.H., A.B. Jaffe, and M. Trajtenberg, *Market Value And Patent Citations.* RAND Journal of Economics, 2005. **36**: p. 16 – 38.

[7] Hall, B.H., A.B. Jaffee, and M. Trajtenberg, *The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools.* National Bureau of Economic Research Working Papers, 2001. **Working Paper Number 8498**.

[8] Kennedy, P., *A guide to Econometrics*. fifth ed. 2003, Cambridge, Massachusettes: MIT Press.

[9] Lanjouw, J.O. and M. Schankerman, *Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators.* The Economic Journal, 2004. **114**: p. 441 - 465.

[10] Lerner, J., *The Importance of Patent Scope: an Empirical Analysis.* RAND Journal of Economics, 1994. **25**(2): p. 319 - 333.

[11] Markman, G.D., M.I. Espina, and P.H. Phan, *Patents as Surrogates for Inimitable and Non-Substitutable Resources.* Journal of Management, 2004. **30**(4): p. 529 – 544.

[12] Nagaoka, S., *Assessing the R&D Management of a Firm in Terms of Speed and Science Linkage: Evidence from US Patents.* Journal of Economics and Management Strategy, 2007. **16**(1): p. 129 - 156.

[13] Nerkar, A. and S. Shane, *Determinants of Invention Commercialization: an Empirical Examination of Academically Sourced Inventions.* Strategic Management Journal, 2007. **28**: p. 1155 - 1166.

[14] Peteraf, M.A., *The Cornerstones of Competitive Advantage: a Resource-Based View.* Strategic Management Journal, 1993. **14**: p. 179 - 191.

[15] Reichert, J.M., *Trends in Development and Approval Times for New Theerapeutics in the United States.* Nature Reviews Drug Discovery, 2003. **2**: p. 695 - 702.

[16] Sorensen, J.B. and T.E. Stuart, *Aging, Obsolescence, and Organizational Innovation* Administrative Science Quarterly, 2000. **45**: p. 81 - 112.